AFRL-RI-RS-TR-2014-292

TEACH OURSELVES: SOCIAL NETWORKS FOR CS STEM EDUCATION

UNIVERSITY OF ARIZONA

DECEMBER 2014

FINAL TECHNICAL REPORT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

STINFO COPY

**AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE**

■ **AIR FORCE MATERIEL COMMAND**　　■ **UNITED STATES AIR FORCE**　　■ **ROME, NY 13441**

NOTICE AND SIGNATURE PAGE

AFRL-RI-RS-TR-2014-292   HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR:

/ S /   / S /
ROBERT L. KAMINSKI   WARREN H. DEBANY JR.
Work Unit Manager   Technical Advisor, Information
                                                         Exploitation and Operations Division
                                                         Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
DEC 2014	FINAL TECHNICAL REPORT	JUN 2010 – JUN 2014

4. TITLE AND SUBTITLE

TEACH OURSELVES: SOCIAL NETWORKS FOR CS STEM EDUCATION

5a. CONTRACT NUMBER
FA8750-10-2-0163

5b. GRANT NUMBER
N/A

5c. PROGRAM ELEMENT NUMBER
61101E

6. AUTHOR(S)

Carole Beal, Paul Cohen

5d. PROJECT NUMBER
BF97

5e. TASK NUMBER
00

5f. WORK UNIT NUMBER
01

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Arizona
888 N. Euclid Ave
Tucson, AZ 85719-4824

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory/RIG
525 Brooks Road
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S)
AFRL/RI

11. SPONSOR/MONITOR'S REPORT NUMBER

AFRL-RI-RS-TR-2014-292

12. DISTRIBUTION AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited. This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Teach Ourselves is a custom-designed and developed online community in which students solve math and science word problems created by other students, and to create and share their own problems with peers. Teach Ourselves includes features that were inspired by recent research on the engaging properties of computer games, including the chance to earn points and badges, to compare progress with other users, and to engage in social activities such as communicating with peers and providing feedback.

15. SUBJECT TERMS

Computer Science, Social Networking, STEM

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON ROBERT KAMINSKI
a. REPORT U	b. ABSTRACT U	c. THIS PAGE U	UU	19	19b. TELEPHONE NUMBER *(Include area code)* N/A

ACKNOWLEDGEMENTS

Table of Contents

1.0 SUMMARY

In an online learning community ("Teach Ourselves"), middle and high school students created, share solved original math and science word problems for points. The overall hypothesis that students could create, share and critique learning objects was supported. Students were able to author math and science problems that fit quality standards, and were able to provide feedback on their peers' work. Results indicated that although students were capable of writing quality problems, they were less adept at evaluating the work of peers. Teachers also allowed some flawed problems created by students to be added into the system, indicating that the task of monitoring students' creative work was highly demanding. Overall, there was high enthusiasm on the part of students and teachers for the activities. Teachers felt that the activities had high value for developing their students' critical thinking skills, as well as their domain knowledge (e.g., understanding of a science topic or math skill). Analysis of log files indicated that many students continued to participate during out-of-school time, suggesting the strong value of the activities for building interest in STEM.

2.0 INTRODUCTION

THE CHALLENGE: INCREASING STEM PARTICIPATION

Over the last decade, new technologies have led to an explosion of user-created content shared and viewed on the Internet, including text, images, videos and even home-authored games. The ease with which content can now be created and shared could help students shift from being *consumers* of educational resources that have been developed by others to *creators* of rich content that can be shared with others.
The Teach Ourselves (TO) project was designed to encourage and engage middle school students with math and science. The middle school years have been identified as a critical point at which many students, particularly girls, lose interest in science, technology, engineering and mathematics (STEM) subjects. Innovative approaches are needed to ensure that students remain engaged with these subjects through high school and beyond.

THE APPROACH

Teach Ourselves is a custom-designed online community in which students solve math and science word problems created by other students, and create and share their own problems with peers. Teach Ourselves includes features that were inspired by recent research on the engaging properties of computer games, including the chance to earn points and badges, to compare progress with other users, and to engage in social activities such as communicating with peers and providing feedback in the form of compliments ("+1") or criticisms (flagging). The number of points that can be earned by solving and creating is determined by a dynamic economy that varies with the number of problems available to solve in various domains. Students can track their points on their profile page, and can compare their performance to others by checking the leaderboards. A more detailed description of the features is provided in METHODS, below.

THEORETICAL FRAMEWORK

In addition to its game-like components, Teach Ourselves is also designed to support creative activity by students. This aspect of the system was inspired by research on the cognitive and motivational benefits of "problem posing." In problem posing, students generate new problems and questions from available information, or seek out information about a topic of interest and use the information to discover new relations (Brown & Walter, 1990; Cai & Huang, 2002; Contreras, 2003; English, 1997; Mestre, 2002; Polya, 1962; Simic-Mullter, Turner & Varley, 2009). Problem posing is thus distinct from the much more common practice of requiring students to solve problems that have been prepared by teachers or that are presented in textbooks. Problem posing is argued to provide students with the opportunity to reflect on what is known

and not known, to restate a problem in a new equivalent form or to vary problems in new ways, and to engage in explanation: all processes that should lead to better problem solving and transfer to new problems (Chi, 2009).

In addition to the hypothesized cognitive benefits, problem posing has also been claimed to increase student motivation, whereas solving problems defined by others day after day often leads to student boredom. Teachers have reported anecdotally that the activity of problem posing leads to class engagement and higher interest, especially among students who are not generally enthusiastic about math and science subjects . Problem posing has also been suggested to help students become more confident and feel a greater sense of "ownership" about the topic (Hausmann & Van Lehn, 2007; King, 1992).

3.0 METHODS

FEATURES OF TEACH OURSELVES

Problem solving
When the student logs in to TO, he or she can view a list of the problems that are already available to be solved (i.e., problems created by prior student users), along with the current points value for each problem. If the student solves a problem within three attempts, he or she earns the points. Each incorrect attempt elicits a brief feedback hint, and the problem solver can also view a multimedia help file created by the problem author. If the student does not enter the correct answer, he or she can try the problem again (although the points value may have fluctuated).

Problem posing
Students can also earn points by creating their own problems. In fact, the values for creating new problems are significantly higher than for solving existing problems, because problem authoring is generally more challenging and time-consuming. To create a new problem, the student works with a template that includes areas for typing in problem text, adding a graphic, entering two pieces of feedback that would be shown if the future problem solver enters incorrect answers, and a help item. Help items can be pictures, slide shows (created with PointPoint), screencast or cell phone videos or other media. Help items are intended to provide an explanation or worked example that can guide the user to the solution but without providing the answer.

When students are ready, they submit their work to their teacher for review. Teachers use an integrated rubric to check that the problem includes accurate and appropriate content, that the answer is correct along with any associated units that need to be specified, and that the attributions for any source materials are listed. If the teacher approves the problem, the student can publish it so that it is available for other students to solve within the TO application, and earns the contracted number of points. Teachers can also return the problem to the author with comments and suggestions for revision.

Social and game-like components
TO includes social networking features such as the ability to +1 ("like"), flag and comment on a problem, along with discussion boards. Also included are game-like features such as leaderboards that show users in terms of overall points, points by domain, class, school and other groupings. Individual progress summaries can be viewed by the student on his or her profile page, including points earned by solving and creating, +1s (compliments) provided by other students, badges earned and stars provided by teachers for high-quality work and helpful feedback given to peers. Students can check their progress and status (badges, compliments, flags) on their account page.

CLASSROOM RESEARCH

The study was conducted with the use of a web-based content management system that allowed students to create their own word problems and to solve problems created by other students. Students earned points for their work, which were displayed in their account pages and on a series of leaderboards. The system was seeded with approximately 100 word problems that had been created by previous users in pilot work so students could immediately begin solving problems. In addition to solving problems, students could create problems in specific domains (applied, earth, life, physical or space sciences, or mathematics). Creating a problem involved writing the problem text, specifying the answer, writing two hints, creating or locating an image, and creating a problem help item which could be a single slide, a slide show, or a video tutorial. The online system provided students with a template for uploading the problem components as they were created. Because creating was relatively labor-intensive, students could earn up to 10 times more points for creating a problem than for solving one.

After the problem was created the student submitted it to the teacher who checked the content of the problem (e.g., appropriate images and content, accurate answer, correct attribution for media, legibility and usefulness of the help item, etc.) using a provided rubric and either approved or denied the problem, providing feedback to the student about what needed to be modified. When the teacher approved a problem the student could publish it, making it available within the system to be solved by peers.

During the activity, some problems created by students and approved by teachers were subsequently censored by the research team and removed from circulation. Some were removed because their format (e.g., multiple-choice, yes-no, true-false) allowed users to solve and earn points simply by trying all possible provided options. The remaining problems were censored because the answer was revealed in some part of the problem (e.g., help or image attribution, in the help text, or the units space); the answer was wrong; the problem was impossible to answer because the author included units in the answer, or lacked specific information for how to submit a correct answer.

As students solved problems created by other students, they could provide feedback about the problem quality through complimenting (+1) a problem or flagging it to communicate issues with the problem (e.g., errors, revealed answers, or inadequate help or attribution). Complimenting and flagging were options but were not required.

Several pilot studies were conducted to provide usability and feasibility data during the development of the Teach Ourselves application. The final evaluation study included 132 students who were 12.3 years old on average; 73 (55%) were girls, and 59 (45%) were boys. Class sizes ranged from 15 to 28 students per teacher. Six classes participated in the evaluation study. Classrooms were located in Tucson, Casa Grande, the greater Phoenix area, and Yuma, AZ. The Teach Ourselves activity ran for approximately 90 days. Students' work with Teach Ourselves was automatically recorded. Both students and teachers participated in online surveys about their experiences with Teach Ourselves.

4.0 RESULTS AND DISCUSSION

The results are presented in terms of the research questions:

DID STUDENTS CREATE QUALITY MATERIALS?

Data were assembled for each student, including the number of points earned from solving and publishing word problems, and the number of problems solved and published. On average, students solved 146 problems and created 5.2 new word problems. Students earned more of their points from solving other students' problems (80% of total points earned) than creating their own (20% of total points earned).

Quality of the problems created by the students was analyzed after the study was completed. A measure of problem quality was established based on a rubric for problem text (i.e., complexity, 0 to 4 possible points), help item quality (i.e., level of helpful information provided, 0 to 4 points), and solvability (i.e., was the problem readily solvable as a contained unit or did it require additional research or learning; 0 to 4 points). The maximum value 12 indicated the highest quality problem. Problems were rated by one of two trained researchers; disagreements based on a subset of 20 problems were rare and were resolved by discussion. On average, students' work was rated 7.5 out of 12 possible points, with a low 4 points to a high of 10 points. Average problem quality varied somewhat across the six classrooms, from a low of 6.6 to a high of 8.5. Not surprisingly, students who earned more points overall also had problems that were independently rated as being of higher quality, suggesting that more extended involvement with TO was associated with better work.

During the Teach Ourselves activities, students created 1,107 original word problems. However, 326 of these problems were subsequently censored for having significant flaws. The research team evaluated the quality of 702 of the remaining problems. (An additional 79 problems were published at the very end of the classroom activity and were not available for the quality analysis.) The mean quality rating provided by researchers was 8.5, with a range of 6 – 11.5 (out of 12 possible points). Not surprisingly, problems created by high school students were rated as higher in quality (M = 9.07) than those created by middle school students (M = 8.33), $F(1,202) = 24.67$, $p < .0001$.

Another indication of quality was provided by students, who had the option to compliment or flag problems created by their peers. Of the 273 students, 55 (20%) both complimented and flagged problems, another 55 (20%) either complimented or flagged a problem, and 163 (60%) never provided either type of feedback to peers. Thus, about 40% of the students provided some feedback about the quality of other students' work. Students flagged 238 problems, and 98 (41.1%) of those flagged problems ended up being censored by researchers due to valid issues with the problem.

Reasons for flagging varied from constructive feedback for the author, or noting a flawed attribution,(e.g., #1969 "Answer is incorrect. Answer should've been O (Oxygen)

instead of H (Hydrogen)", #3208 "Not enough clarification on how to correctly answer the question." #464 : "Your help item gives the problem away.", #2763 : "The help item isn't helpful and irrelenvent (sic)!"). However, 140 (58.8%) problems were flagged by students but remained in circulation after the researchers reviewed them and found no valid issues. These cases appeared to involve the problem solver making a mistake (i.e., misinterpreting the problem) that led him or her to assume the problem was incorrect.

With regard to compliments, 723 problems were given a +1 by a student who had solved it (note that each problem could be solved by multiple students). Interestingly, 219 (30.3%) of these problems were subsequently censored because they were seriously flawed. In fact, 112 of these problems (i.e., complimented and later censored) received compliments from three or more different students.

HOW DID STUDENTS RESPOND TO TEACH OURSELVES?

Independent learning
Survey results indicated that students liked the activity, and their comments pointed to the value of the approach in terms of students' being able to direct their own learning. Sample comments included the following (spelling and grammar errors in original comments are retained here):
- It's very independent and I like that
- I like that the whole thing were made by students
- It's fun that all the questions are made by kids like us

Game elements
The game-like elements (e.g., points, badges and leaderboards) in Teach Ourselves were especially appreciated by students. Sample comments included:
- Mrs. Brewer, Guess what! Were in top 10 for groups in the lader board!!! I saw it said Canyon Ridge! ☺ PS were doing great!
- I like that you can get points and badges
- Getting points motivates me!
- One thing I really like about Teach Ourselves is the idea of competition. The leader boards really keep kids on their toes to try and get to the top
- I trying to git in first ranking.
- I love Teach Ourselves. I like the thrill of getting points and getting on the leader board."

Out-of-school participation
Another indication of the positive reaction by students was that they did not limit their participation to class hours. Navigation events logged for each student were extracted and the event timestamps were used to categorize the events by day (weekday or weekend; holiday weekdays were counted as weekend days) and time (by hour, from midnight to midnight). School activity was defined as events occurring between 7:30 a.m. and 3:00 p.m. Evening activity was defined as actions occurring between 8:00

p.m. and midnight. The total number of navigation events in these categories was calculated for each student.

Students had an average of 1,210 total navigation events recorded during the study, with a range from 50 to 7,318. Of these, 49% occurred during out-of-school hours, meaning at times other than Mondays-Fridays from 7:30 a.m. to 3:00 p.m. In fact, only 27 students (20% of the sample) *never* accessed TO during out-of-school hours. Looking more closely at the out-of-school access information, most appeared to be in the evening hours (8:00 p.m. through 12 midnight); evening use accounted for about 15% of the total navigation events. Weekend activity accounted for about 9% of the navigation events, with 35% of the students logging in at least once on a weekend day. Interestingly, there was a significant correlation between the number of events during evening hours and events on weekend days, suggesting that those students who were more engaged with TO after school were also likely to check in over the weekends.

The average percentage of navigations that occurred out of school was compared across the six classrooms. The results of a one-way ANOVA indicated that there was significant variation across the classrooms, $F(5,126) = 34,195$, $p < .01$. One teacher reported that she used TO primarily as a homework activity, so the high percentage (90%) of events occurring out of school for her students was not unexpected. A second classroom included students who rarely (11.6%) accessed the system outside of school. This left four classrooms (85 students) where TO was implemented by teachers during the school day. In these classrooms, 45% of the navigation events still occurred outside of school hours.

HOW DID TEACHERS RESPOND TO TEACH OURSELVES?

Teachers were highly enthusiastic about the Teach Ourselves activities. Even though they reported that it was sometimes challenging to keep up with the reviewing process, their comments were generally positive, especially with regard to the impact of the activity on students' critical thinking skills. Teachers identified many benefits to using TO with their students including self-evaluation, critical thinking, digital literacy, and reinforcing STEM knowledge. All (100%) said that they thought their students had enjoyed TO and that it had helped them learn domain-specific material; 89% said it helped students improve higher-order thinking. Sample comments included the following:

- Students are using higher-level thinking by critiquing and questioning each other
- Exciting program – my students use the knowledge we gain in class and apply it to a real world problem
- I really think this is helping my students with their higher-order thinking
- I'm amazed at how creative and well-written some of my students' questions are
- My students have already been asking if they can access it in the summer
- Businesses are seeking out students who can analyze a problem and find a way to solve it. TO is a wonderful platform on which students can build these skills

- DEFINITELY interested in using it next year!
- Their excitement about the program was evident - they LOVED having Teach Ourselves days.
- I witnessed my students take pride on their work and become more confident and sure of themselves.

5.0 CONCLUSIONS

Overall, about 70% of the word problems authored by students met the standards for being clear, solvable and helpful (in that hints and scaffolding were included). Thus, one conclusion from the project was that students were able to create quality word problems in math and science domains given a) the availability of technology designed to structure the activity and b) given the support of teachers who were required to review students' work and provide feedback. Although the role of technology as a facilitator was not directly assessed here, it seemed apparent that the ability to search for interesting facts around which to build problems, to find images online, to use digital tools to create help items, and to share the finished work within the online repository made the activity more feasible than would have been the case with, say, paper-and-pencil activities. The gamified elements of earning points, badges and leaderboard rankings were also made practical by the technology platform, and these elements were clearly important to students. The finding that 80% of the students used TO out of school was striking given the focus on math and science topics, which are not always highly appealing to many middle and high school students. Adolescents spend a great deal of their out-of-school time engaged with entertainment media and social networking. The experience with Teach Ourselves suggests that incorporating some of the social and game-like elements into an academic application may have the potential to bridge formal and informal learning.

One critical factor is the involvement of teachers in acting as the first line of quality control. The finding that about 30% of the student-created problems were ultimately removed due to fatal flaws is an indication of the demanding nature of problem posing activities for teachers. On average, each teacher had more than 100 unique word problems to review and approve. Given the volume, perhaps it is not surprising that not all problem errors were caught. Comments from some teachers also suggested that on occasion they intentionally approved a flawed problem because they did not want to undermine a student's motivation; this warrants additional investigation.

Additional findings involved students' feedback to peers about the quality of their work. Only about 40% of the students utilized the feedback features in the system, which were optional. Interestingly, when students did compliment their peers' work, they often did so on problems that included fatal flaws. It seems likely that this behavior was socially driven, for example, giving +1 to friends and classmates, regardless of the actual problem quality. Students were also not terribly effective at identifying poor quality problems with the flagging feature. Fewer than half of the flagged problems were ultimately censored and, in many cases, it appeared that the problem itself was adequate but that the flagger had missed critical information when trying to solve it.

In summary, the study results demonstrated that students were able to create problems that met the standards for being solvable and including appropriate scaffolding. However, there were indications that the activity was demanding: First, students' participation, although enthusiastic (as reported by teachers and by out-of-school participation), was somewhat constrained. About 25% of the students avoided creating

any of their own problems even though they could earn more points by doing so than by solving. About 60% did not provide any quality feedback to their peers via the +1 and flagging options. When students did compliment or flag, their feedback was not always related to the actual problem quality. Second, although teachers were initial checkpoint for problem quality, they did not always catch or act on poor-quality problems. In fact, about a third of the problems that were created by students and approved by teachers were censored for having fatal flaws. Additional research is needed to understand more clearly how best to support teachers in implementing problem posing activities, as well as to identify best practices for obtaining the benefits for learning that are outlined in the theoretical framework associated with problem posing.

The primary limitation of the research was the need to position the activities in the context of the traditional face-to-face classroom. The original vision of an online community directed by students via social media (a Facebook app) had to be entirely reconceptualized into a classroom-based activity, due to after-award restrictions imposed by the Department of Defense requirement that students have written parent consent for participation. This was disappointing given that all activities, including those proposed to be based in Facebook, had been reviewed and approved by the University of Arizona Institutional Review Board. When the project was implemented in schools, the middle and high school administrators saw value in and were supportive of the project activities, as were parents who provided near-universal consent for their child's data to be used in the research. The project experience provides an indication that Department of Defense standards for research approval may need to be re-visited in light of the new opportunities for participation in virtual spaces.

6.0 REFERENCES

Brown, S. I., & Walter, M. I. (1990). The art of problem posing. Hillsdale NJ: Erlbaum.

Cai, J., & Huang, S. (2002). Generalized and generative thinking in U.S. and Chinese students' mathematical problem solving and problem posing. Journal of Mathematical Behavior, 21, 401-421.

Polya, G. (1962). Mathematical discovery: On understanding, learning, and teaching problem solving, Vol. 1. New York: John Wiley.

Chi, M.T.H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. Topics in Cognitive Science,1, 73-105.

Contreras, J. N. (2003). A problem-posing approach to specializing, generalizing and extending problems with interactive geometry software. The Mathematics Teacher, 96, 270-275.

English, L. (1997, Nov.). Promoting a problem-posing classroom. Teaching Children Mathematics, 172-179.

Hausmann, R., & Van Lehn, K. (2007). Explaining self-explaining: A contrast between content and generation. Proceedings of the 13th International Conference on Artificial Intelligence and Education, IOS Press.

King, A. (1992). Comparison of self-questioning, summarizing and note-taking review as strategies for learning from lectures. American Educational Research Journal, 29, 303-323.

Mestre, J. P. (2002, Jan.). Probing adults' conceptual understanding and transfer of learning via problem posing. Journal of Applied Developmental Psychology, 23, 9-50.

Simic-Mullter, K., Turner, E., & Varley, M. C. (2009). Math Club problem posing. Teaching Children Mathematics, 16, 206-212.

APPENDIX - PUBLICATIONS

Beal, C. R., & Cohen, P. R. (2012). Teach Ourselves: Technology for problem posing in the STEM classroom. Creative Education, 3, 513-519.

Beal, C. R., Strohm, J., Schwindy, L., & Cohen, P. R. (2013). Teach Ourselves: A peer-to-peer learning community linking in- and out-of-class activity. Bulletin of the IEEE Technical Committee on Learning Technology, 15, 13-16.

LIST OF ACRONOYMS

AZ	Arizona
H	Hydrogen
O	Oxygen
STEM	Science, Technology, Engineering and Mathematics
TO	Teach Ourselves